

RECEIVED:



TECH CENTER 1600-2900

OIPE

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/083,768

DATE: 03/27/2002

TIME: 13:24:50

Input Set A:\PK3065US3 SeqList.txt

Output Set: N:\CRF3\03272002\J083768.raw

## SEQUENCE LISTING

## 4 (1) GENERAL INFORMATION.

6 (i) APPLICANT: Dower, William J.  
 7 Barrett, Ronald W.  
 8 Cwirla, Steven E.  
 9 Duffin, David J.  
 10 Gates, Christian  
 11 Haseiden, Sherril S.  
 12 Mattheakis, Larry C.  
 13 Schatz, Peter J.  
 14 Wagstrom, Christopher R.  
 15 Wrighton, Nicholas C.

17 (ii) TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
 18 THROMBOPOEITIN RECEPTOR

20 (iii) NUMBER OF SEQUENCES: 232

22 (iv) CORRESPONDENCE ADDRESS:

23 (A) ADDRESSEE: Glaxo Wellcome  
 24 (B) STREET: Five Moore Drive, P.O. Box 13398  
 25 (C) CITY: Research Triangle Park  
 26 (D) STATE: NC  
 27 (E) COUNTRY: USA  
 28 (F) ZIP: 27709

30 (v) COMPUTER READABLE FORM:

31 (A) MEDIUM TYPE: Floppy disk  
 32 (B) COMPUTER: IBM PC compatible  
 33 (C) OPERATING SYSTEM: PC-DOS/MS-DOS  
 34 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

35 (vi) CURRENT APPLICATION DATA:

37 (A) APPLICATION NUMBER: US/10/083,768  
 38 (B) FILING DATE: 27-Feb-2002

40 (viii) ATTORNEY/AGENT INFORMATION:

41 (A) NAME: Hrubiec, Robert T  
 42 (B) REGISTRATION NUMBER: 36 392  
 43 (C) REFERENCE/DOCKET NUMBER PK3065USW

45 (ix) TELECOMMUNICATION INFORMATION:

46 (A) TELEPHONE: 919-248-1000

49 (2) INFORMATION FOR SEQ ID NO: 1

51 (i) SEQUENCE CHARACTERISTICS:

52 (A) LENGTH: 5 amino acids  
 53 (B) TYPE: amino acid  
 54 (C) STRANDEDNESS  
 55 (D) TOPOLOGY linear

57 (ii) MOLECULE TYPE: peptide

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/083,768

DATE: 03/27/2002  
TIME: 13:24:50

Input Set : A:\PK3065US3 SeqList.txt  
Output Set: N:\CRF3\03272002\J083768.raw

52 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
W--> 64 Trp Ser Xaa Trp Ser  
65 1 5  
67 (2) INFORMATION FOR SEQ ID NO: 2:  
69 (i) SEQUENCE CHARACTERISTICS:  
70 (A) LENGTH: 7 amino acids  
71 (B) TYPE: amino acid  
72 (C) STRANDEDNESS:  
73 (D) TOPOLOGY: linear  
75 (ii) MOLECULE TYPE: peptide  
80 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
W--> 82 Xaa Xaa Xaa Xaa Xaa Xaa Xaa  
83 1 5  
85 (2) INFORMATION FOR SEQ ID NO: 3:  
87 (i) SEQUENCE CHARACTERISTICS:  
88 (A) LENGTH: 9 amino acids  
89 (B) TYPE: amino acid  
90 (C) STRANDEDNESS:  
91 (D) TOPOLOGY: linear  
93 (ii) MOLECULE TYPE: peptide  
98 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
W--> 100 Xaa Gly Xaa Xaa Xaa Xaa Trp Xaa  
101 1 5  
103 (2) INFORMATION FOR SEQ ID NO: 4:  
105 (i) SEQUENCE CHARACTERISTICS:  
106 (A) LENGTH: 9 amino acids  
107 (B) TYPE: amino acid  
108 (C) STRANDEDNESS:  
109 (D) TOPOLOGY: linear  
111 (ii) MOLECULE TYPE: peptide  
116 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
W--> 118 Xaa Gly Pro Thr Leu Arg Xaa Trp Xaa  
119 1 5  
121 (2) INFORMATION FOR SEQ ID NO: 5:  
123 (i) SEQUENCE CHARACTERISTICS:  
124 (A) LENGTH: 10 amino acids  
125 (B) TYPE: amino acid  
126 (C) STRANDEDNESS:  
127 (D) TOPOLOGY: linear  
129 (ii) MOLECULE TYPE: peptide  
134 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
W--> 136 Xaa Xaa Gly Xaa Xaa Xaa Xaa Trp Xaa  
137 1 5 10  
139 (2) INFORMATION FOR SEQ ID NO: 6:  
141 (i) SEQUENCE CHARACTERISTICS:  
142 (A) LENGTH: 18 amino acids  
143 (B) TYPE: amino acid  
144 (C) STRANDEDNESS:  
145 (D) TOPOLOGY: linear

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/083,768

DATE: 03/27/2002  
TIME: 13:24:59

Input Set : A:\PK3065US3 SeqList.txt  
Output Set: N:\CRF3\03272002\J083768.raw

147 (iii) MOLECULE TYPE: peptide  
152 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
154 Gly Gly Cys Ala Asp Gly Pro Thr Leu Arg Gln Trp Ile Ser Phe Cys  
155 1 5 10 15  
157 Gly Gly  
160 (2) INFORMATION FOR SEQ ID NO: 7:  
162 (i) SEQUENCE CHARACTERISTICS:  
163 (A) LENGTH: 19 amino acids  
164 (B) TYPE: amino acid  
165 (C) STRANDEDNESS:  
166 (D) TOPOLOGY: linear  
168 (ii) MOLECULE TYPE: peptide  
173 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7  
175 Gly Asn Ala Asp Gly Pro Thr Leu Arg Gln Trp Leu Glu Gly Arg Arg  
176 1 5 10 15  
178 Pro Lys Asn  
181 (2) INFORMATION FOR SEQ ID NO: 8:  
183 (i) SEQUENCE CHARACTERISTICS:  
184 (A) LENGTH: 19 amino acids  
185 (B) TYPE: amino acid  
186 (C) STRANDEDNESS:  
187 (D) TOPOLOGY: linear  
189 (ii) MOLECULE TYPE: peptide  
194 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
196 Gly Gly Cys Ala Asp Gly Pro Thr Leu Arg Glu Trp Ile Ser Phe Cys  
197 1 5 10 15  
199 Gly Gly Lys  
202 (2) INFORMATION FOR SEQ ID NO: 9:  
204 (i) SEQUENCE CHARACTERISTICS:  
205 (A) LENGTH: 18 amino acids  
206 (B) TYPE: amino acid  
207 (C) STRANDEDNESS:  
208 (D) TOPOLOGY: linear  
210 (ii) MOLECULE TYPE: peptide  
215 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9  
217 Thr Ile Lys Gly Pro Thr Leu Arg Gln Trp Leu Lys Ser Arg Glu His  
218 1 5 10 15  
220 Thr Ser  
223 (2) INFORMATION FOR SEQ ID NO: 10:  
225 (i) SEQUENCE CHARACTERISTICS:  
226 (A) LENGTH: 18 amino acids  
227 (B) TYPE: amino acid  
228 (C) STRANDEDNESS:  
229 (D) TOPOLOGY: linear  
231 (ii) MOLECULE TYPE: peptide  
236 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
238 Ser Ile Glu Gly Pro Thr Leu Arg Glu Trp Leu Thr Ser Arg Thr Pro  
239 1 5 10 15  
241 His Ser

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/083,768

DATE: 03/27/2002  
TIME: 13:24:50

Input Set : A:\PK3065US3 SeqList.txt  
Output Set: N:\CRF3\03272002\J083768.raw

244 (2) INFORMATION FOR SEQ ID NO: 11:  
246 (i) SEQUENCE CHARACTERISTICS:  
247 (A) LENGTH: 19 amino acids  
248 (B) TYPE: amino acid  
249 (C) STRANDEDNESS:  
250 (D) TOPOLOGY: linear  
252 (ii) MOLECULE TYPE: peptide  
257 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:  
259 Leu Ala Ile Glu Gly Pro Thr Leu Arg Gln Trp Leu His Gly Asn Gly  
260 1 5 10 15  
262 Arg Asp Thr  
265 (2) INFORMATION FOR SEQ ID NO: 12:  
267 (i) SEQUENCE CHARACTERISTICS:  
268 (A) LENGTH: 14 amino acids  
269 (B) TYPE: amino acid  
270 (C) STRANDEDNESS:  
271 (D) TOPOLOGY: linear  
273 (ii) MOLECULE TYPE: peptide  
278 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:  
280 Cys Ala Asp Gly Pro Thr Leu Arg Glu Trp Ile Ser Phe Cys  
281 1 5 10  
283 (2) INFORMATION FOR SEQ ID NO: 13:  
285 (i) SEQUENCE CHARACTERISTICS:  
286 (A) LENGTH: 14 amino acids  
287 (B) TYPE: amino acid  
288 (C) STRANDEDNESS:  
289 (D) TOPOLOGY: linear  
291 (ii) MOLECULE TYPE: peptide  
296 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:  
298 Ile Glu Gly Pro Thr Leu Arg Gln Trp Leu Ala Ala Arg Ala  
299 1 5 10  
301 (2) INFORMATION FOR SEQ ID NO: 14:  
303 (i) SEQUENCE CHARACTERISTICS:  
304 (A) LENGTH: 7 amino acids  
305 (B) TYPE: amino acid  
306 (C) STRANDEDNESS:  
307 (D) TOPOLOGY: linear  
309 (ii) MOLECULE TYPE: peptide  
314 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:  
W--> 316 Cys Xaa Xaa Xaa Xaa Xaa Xaa  
317 1 5  
319 (2) INFORMATION FOR SEQ ID NO: 15:  
321 (i) SEQUENCE CHARACTERISTICS:  
322 (A) LENGTH: 16 amino acids  
323 (B) TYPE: amino acid  
324 (C) STRANDEDNESS:  
325 (D) TOPOLOGY: linear  
327 (ii) MOLECULE TYPE: peptide  
332 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

RAW SEQUENCE LISTING  
PATENT APPLICATION: US/10/083,768

DATE: 03/27/2002  
TIME: 13:24:50

Input Set : A:\PK3065US3 SeqList.txt  
Output Set N:\CRF3\03272002\J083768.raw

334 Gly Gly Cys Ile Leu Arg Gln Ile Leu His Gly Gly Phe Cys Ile Ser  
335 1 5 10 15

338 (2) INFORMATION FOR SEQ ID NO: 16:

340 (i) SEQUENCE CHARACTERISTICS:

341 (A) LENGTH: 8 amino acids

342 (B) TYPE: amino acid

343 (C) STRANDEDNESS:

344 (D) TOPOLOGY: linear

346 (ii) MOLECULE TYPE: peptide

351 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

W--> 353 Xaa Cys Xaa Xaa Xaa Xaa Xaa Xaa

354 1 5

356 (2) INFORMATION FOR SEQ ID NO 17:

358 (i) SEQUENCE CHARACTERISTICS:

359 (A) LENGTH: 15 amino acids

360 (B) TYPE: amino acid

361 (C) STRANDEDNESS:

362 (D) TOPOLOGY: linear

364 (ii) MOLECULE TYPE: peptide

369 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 17:

371 Ile Glu Gly Pro Thr Leu Arg Gln Trp Leu Ala Ala Arg Ala Lys

372 1 5 10 15

374 (2) INFORMATION FOR SEQ ID NO: 18:

376 (i) SEQUENCE CHARACTERISTICS:

377 (A) LENGTH: 16 amino acids

378 (B) TYPE: amino acid

379 (C) STRANDEDNESS:

380 (D) TOPOLOGY: linear

382 (ii) MOLECULE TYPE: peptide

385 (ix) FEATURE:

386 (A) NAME/KEY: Modified-site

387 (B) LOCATION: 15

388 (D) OTHER INFORMATION: /product= "Beta-ala"

391 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

W--> 393 Ile Glu Gly Pro Thr Leu Arg Gln Trp Leu Ala Ala Arg Ala Xaa Lys

394 1 5 10 15

397 (2) INFORMATION FOR SEQ ID NO: 19:

399 (i) SEQUENCE CHARACTERISTICS:

400 (A) LENGTH: 93 base pairs

401 (B) TYPE: nucleic acid

402 (C) STRANDEDNESS: double

403 (D) TOPOLOGY: linear

405 (ii) MOLECULE TYPE: cDNA

408 (ix) FEATURE:

409 (A) NAME/KEY: CDS

410 (B) LOCATION: 1..84

413 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

415 CTC GAG AGC GGG TAA GTG GTG CAT GGG GAG CAG GTG GGT GGT GAG GCC

416 Leu Glu Ser Gly Gln Val Val His Gly Glu Gln Val Gly Glu Ala

48

VERIFICATION SUMMARY  
PATENT APPLICATION: US/10/083,768

DATE: 03/27/2002  
TIME: 13:24:51

Input Set : A:\PK3065US3 SeqList.txt  
Output Set: N:\CRF3\03272002\J083768.raw

L:37 M:220 C: Keyword misspelled or invalid format. ((A) APPLICATION NUMBER:  
L:38 M:220 C: Keyword misspelled or invalid format. ((B) FILING DATE:  
L:64 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1  
L:82 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2  
L:100 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3  
L:118 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4  
L:136 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5  
L:316 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:14  
L:353 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:16  
L:393 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:18  
L:457 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21  
L:460 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:21  
L:478 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22  
L:481 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22  
L:499 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23  
L:502 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:23  
L:520 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24  
L:523 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:24  
L:1651 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:84  
L:1664 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:85  
L:2812 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:147  
L:2854 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:149  
L:3322 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:172  
L:3362 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:174  
L:3389 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:175  
L:3411 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:176  
L:3438 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:177  
L:3465 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:178  
L:3487 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:179  
L:3509 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:180  
L:3530 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:184  
L:3814 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:196  
L:3854 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:198  
L:3894 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:200  
L:3934 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:202  
L:3974 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:204  
L:4327 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:223  
L:4345 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:224  
L:4418 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:227  
L:4433 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:228  
L:4503 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:231  
L:4523 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:232